GIS Basics

SPATIAL STATISTICS WORKSHOP

March 15 – 16, 2006

Presenter: D. Alan Davenport, GIS Coordinator
Division of Migratory Bird Management
U. S. Fish and Wildlife Service, Laurel, MD

Acknowledgement:

Some of the slides used in this presentation were adapted from the course **TEC7112 – *GIS Introduction for Conservation Professionals*** taught at the National Conservation Training Center, Sheperdstown, WV.

What is a Geographic Information System?

A GIS is

- A computer-based system designed for the collection, storage, and analysis of phenomena where geographic (spatial) location is an important characteristic or critical to the analysis.

Components

- Spatial Data
- Attributes

Spatial Data

- Landscape elements that have physical dimensions and geographical location. These elements can be represented in two different ways:
- Vectors
- Rasters

Vector Data

Points

Lines

Polygons

Vector Data

Points

- locations of buildings
- wood duck boxes
- water control structures

Vector Data

Lines

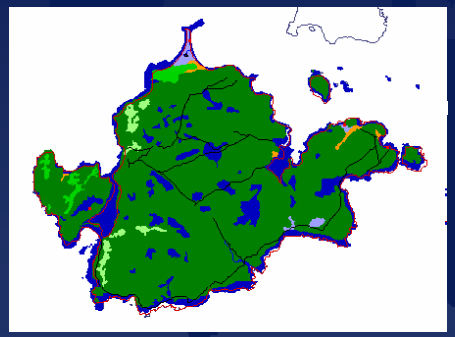
- roads
- boundaries
- streams
- power lines

Vector Data

Polygons

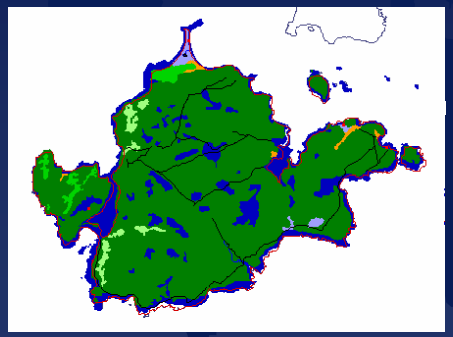
- lakes
- cover types
- timber stands

Vector Data



Overview

Vector Data



Overview

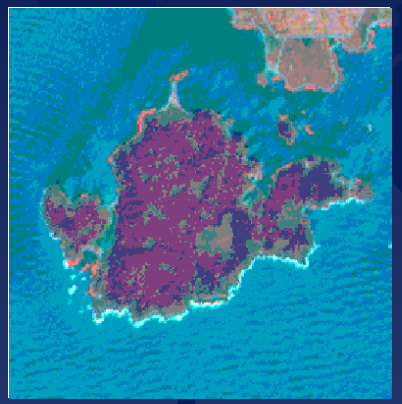
Close-up

Raster Data

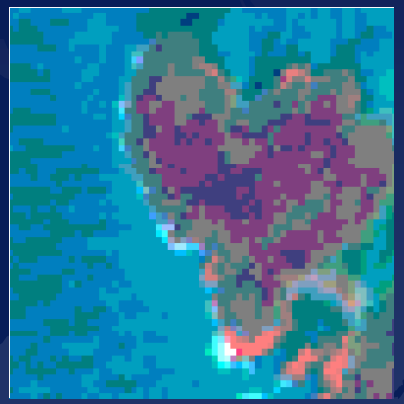
Cells or Pixels

- Landscape elements represented as rows and columns of continuous cells
- Each cell has a location
- Each cell has a value or attribute

Raster Data



Overview



Close-up

Raster Data

- Considerations:

Each cell is a rectangle or square of a constant size. The size of the cells determines the resolution of the map. As the cell size decreases the map resolution increases, but so does the storage requirement in the computer.

Raster Data

Overview

Attributes

- The number of eggs in wood duck box number 27.
- The level of water at Lake Sepik on 27 June 1994.
- The name of a road.
- The volume of red oak saw logs in timber stand number 4.
- The number of black duck broods in Hayes Flowage in 1994.

How do we put it all together?

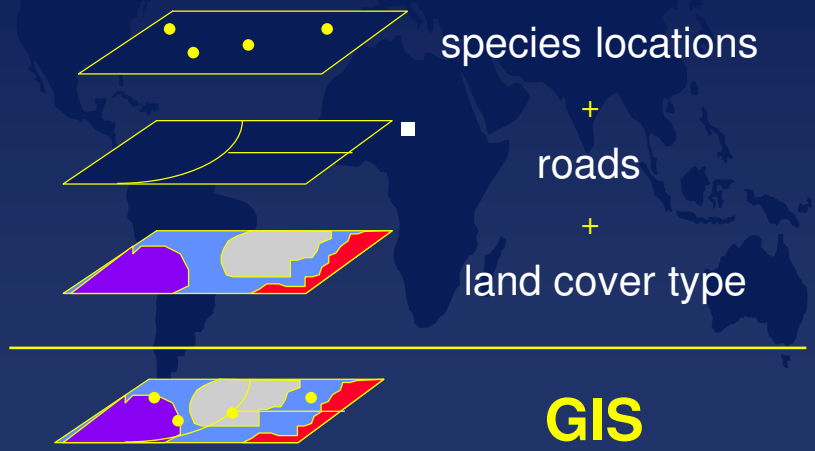
To use spatial data in a GIS you need to know:

- Where each feature is located ***(Coordinates)***
  Geographical Coordinates, X and Y
- What each feature represents ***(Attributes)***
  Can be any number of descriptive characteristics, but there must be at least one.
- Relationships among features ***(Topology)***
  The logic that connects the features to each other, for example, how the location of a wood duck box relates to the location of the nearest wetland. Topology is internally managed by the GIS software.

Spatial data and its attributes must be arranged in a logical order to create a GIS.

This arrangement is a series of layers, or ***THEMES***, each which share a common coordinate system.

A GIS consists of *Data Layers or Themes*



species locations

+

roads

+

land cover type

GIS

The ultimate purpose of a GIS is to answer spatial questions...

...NOT necessarily to make 'PRETTY' maps!

Typical questions include:

- What is at .....?
- Where is .....?
- What has changed since .....?
- What spatial patterns exist?
- What if .....?

An important thing to remember...

The questions must be asked before the data are developed.

GIS Software

http://esri.com/

http://imgs.intergraph.com/

http://www.mapinfo.com/location/integration

http://www.genaware.com/products/genamap/

Who is ESRI ?

•Environmental Systems Research Institute, Redlands, CA

http://www.esri.com/index.html

ArcGis

What is ArcGIS?

- *An integrated collection of GIS software products* for building a complete GIS. The ArcGIS framework enables you to deploy GIS functionality—in desktops, servers (including the Web), or mobile devices

Why ArcGIS?

- The defacto GIS software standard within the FWS

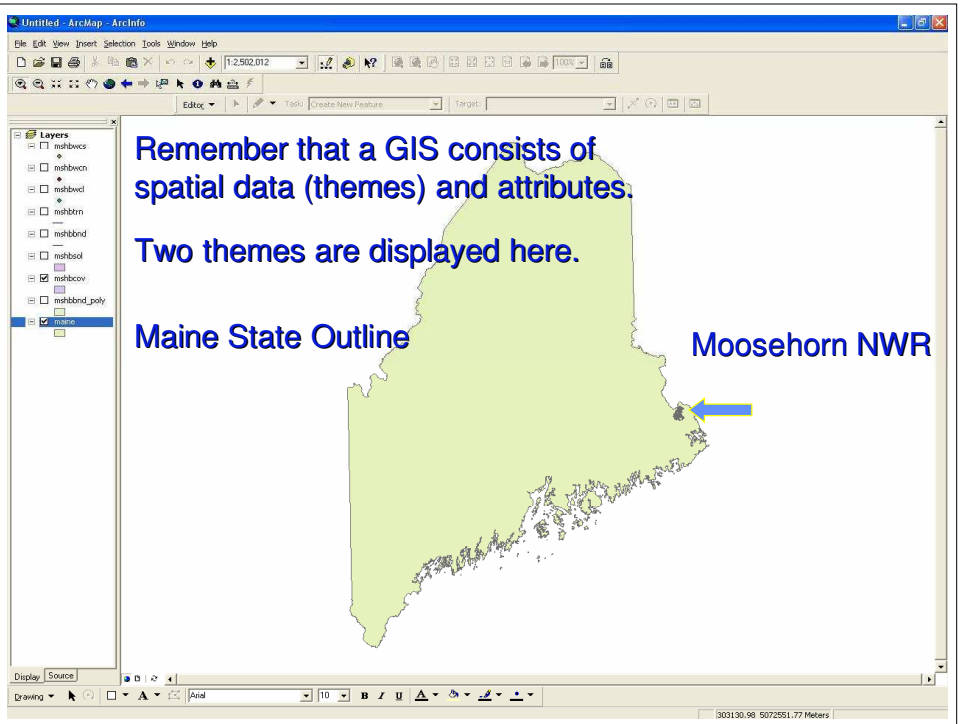
Who else uses ArcGIS?

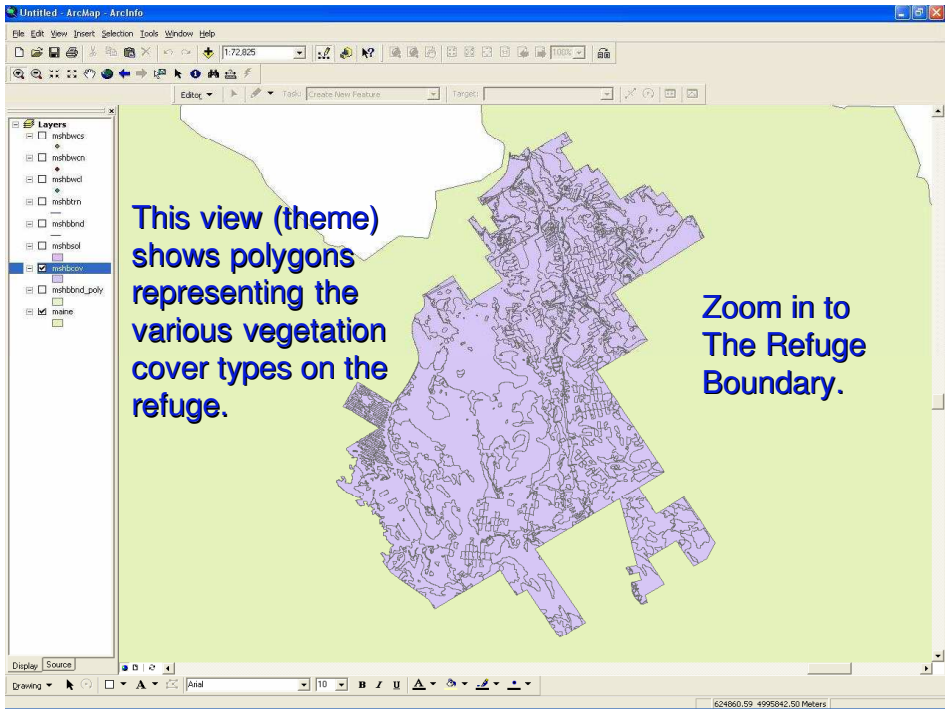
- Most Federal & State Land Management Agencies
- USGS, Forest Service, NPS, BLM, FWS

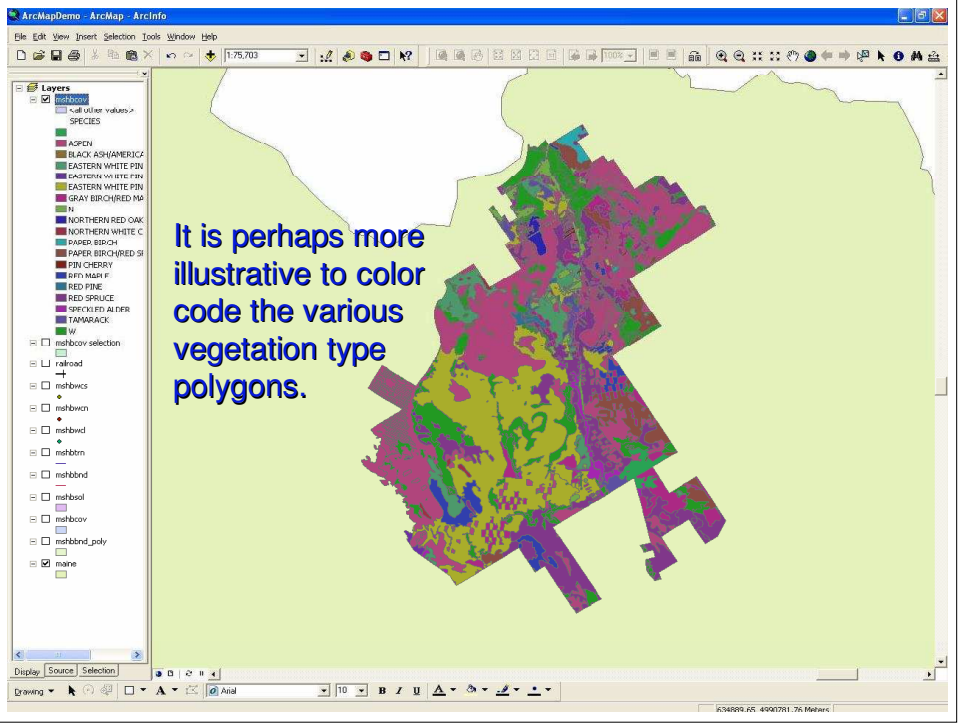
Spatial Analysis Exercises

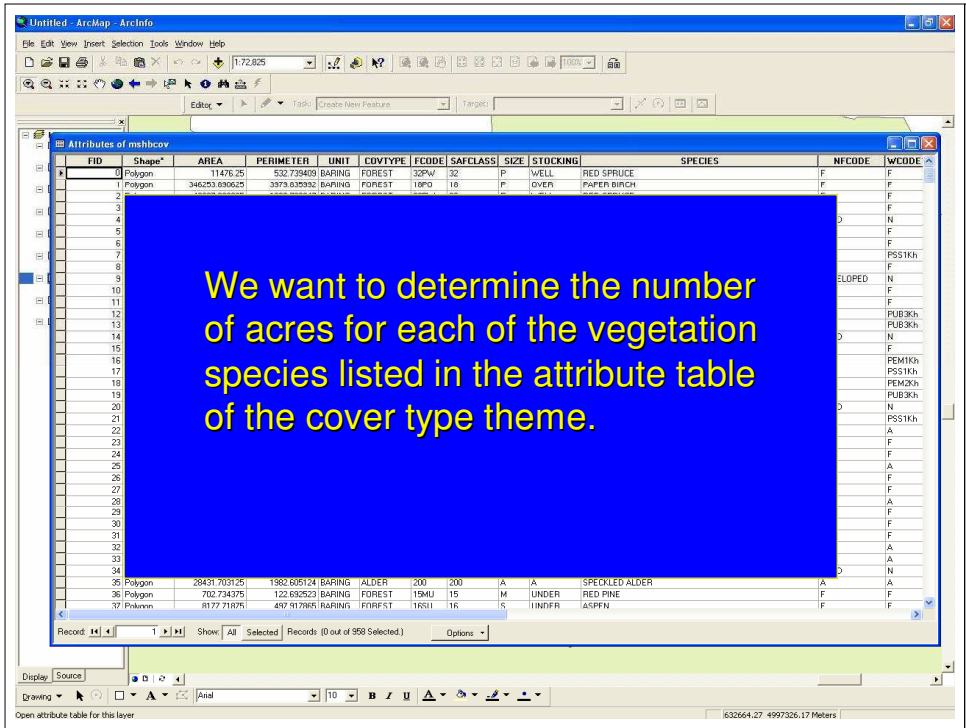
Using ArcMap

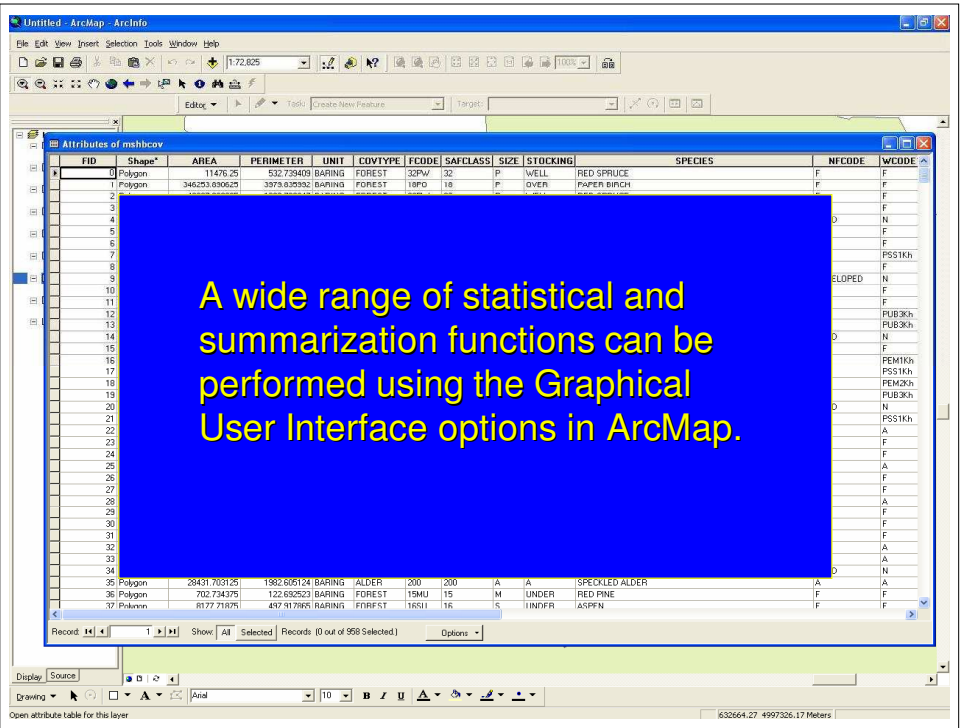
- Perform simple Descriptive Statistical Analyses
- Conduct Complex Spatial Analyses

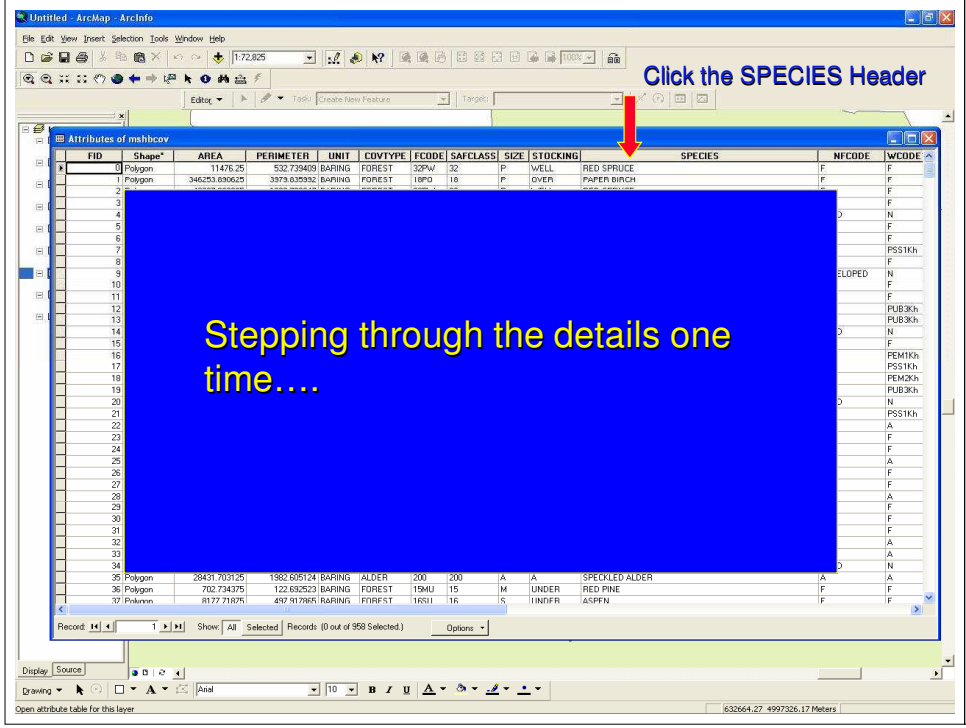


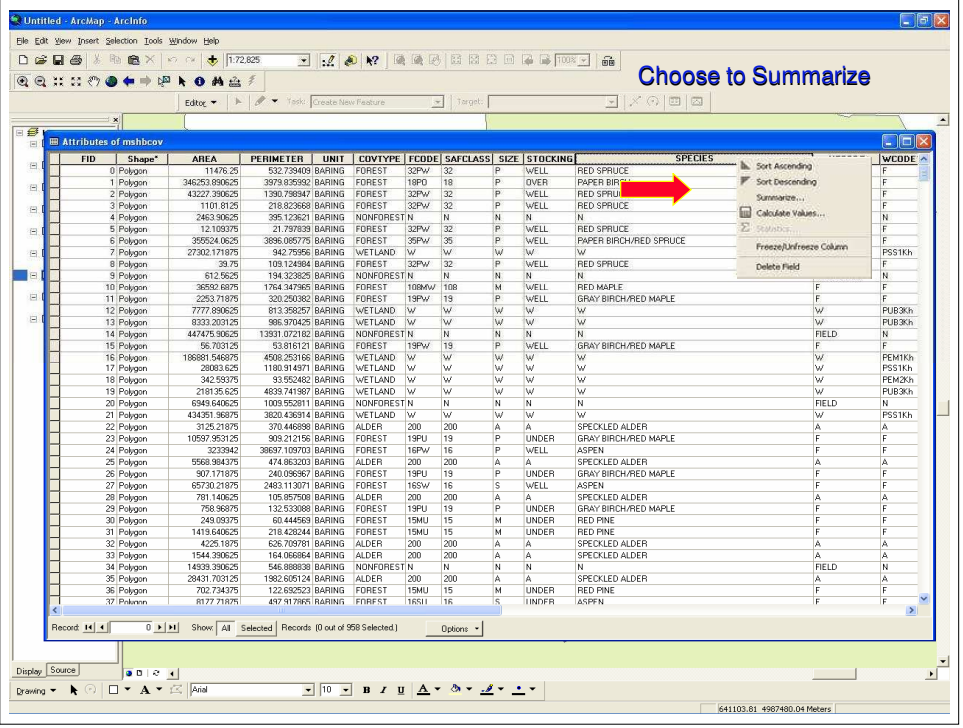


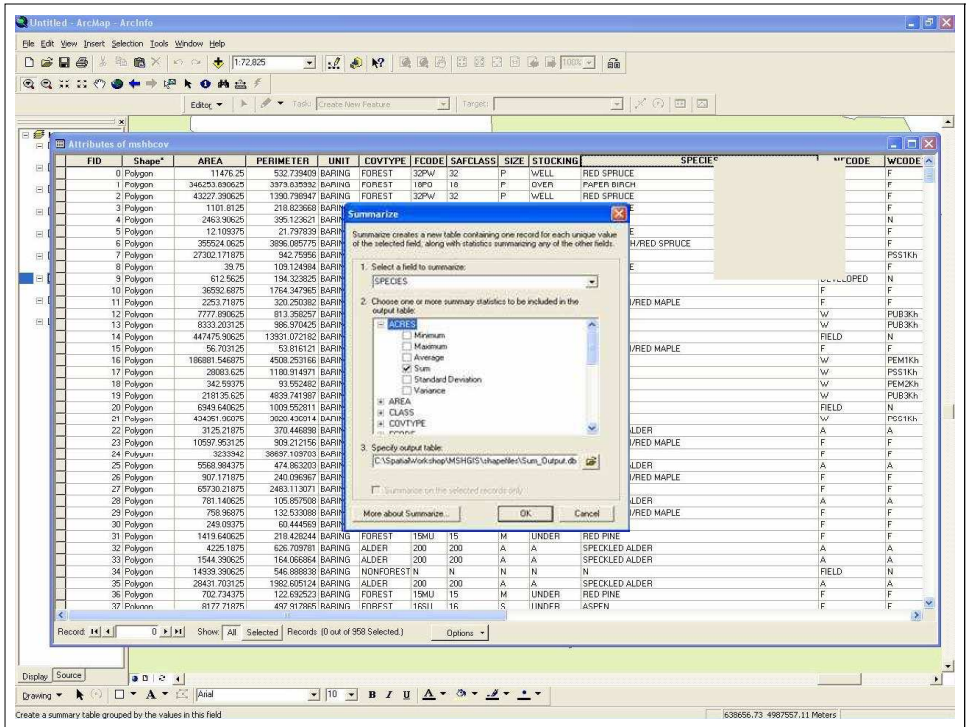


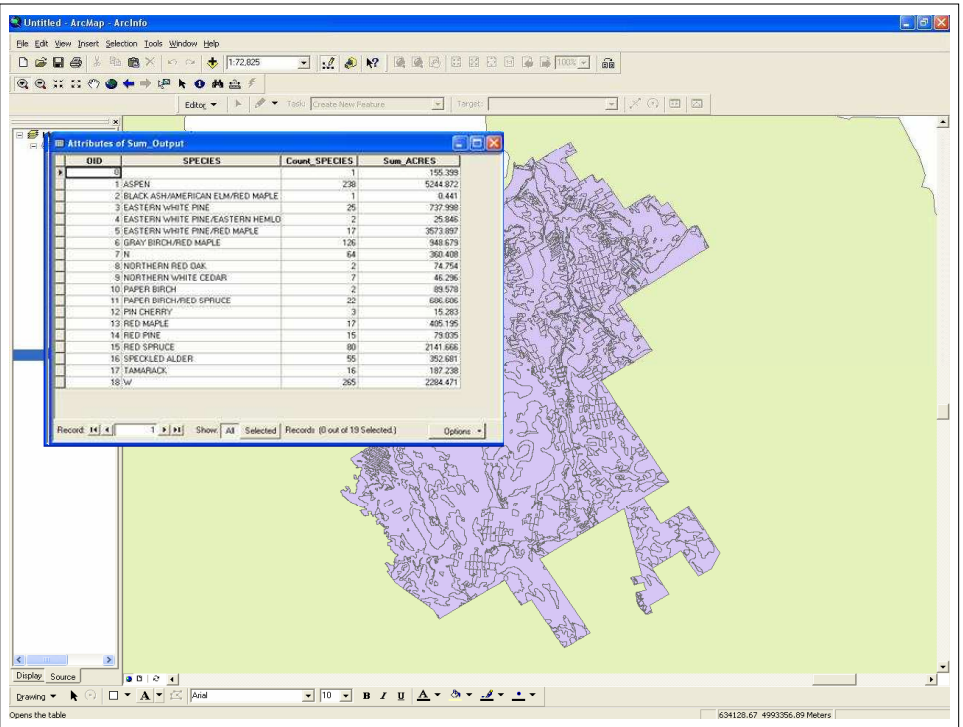

We want to determine the number of acres for each of the vegetation species listed in the attribute table of the cover type theme.

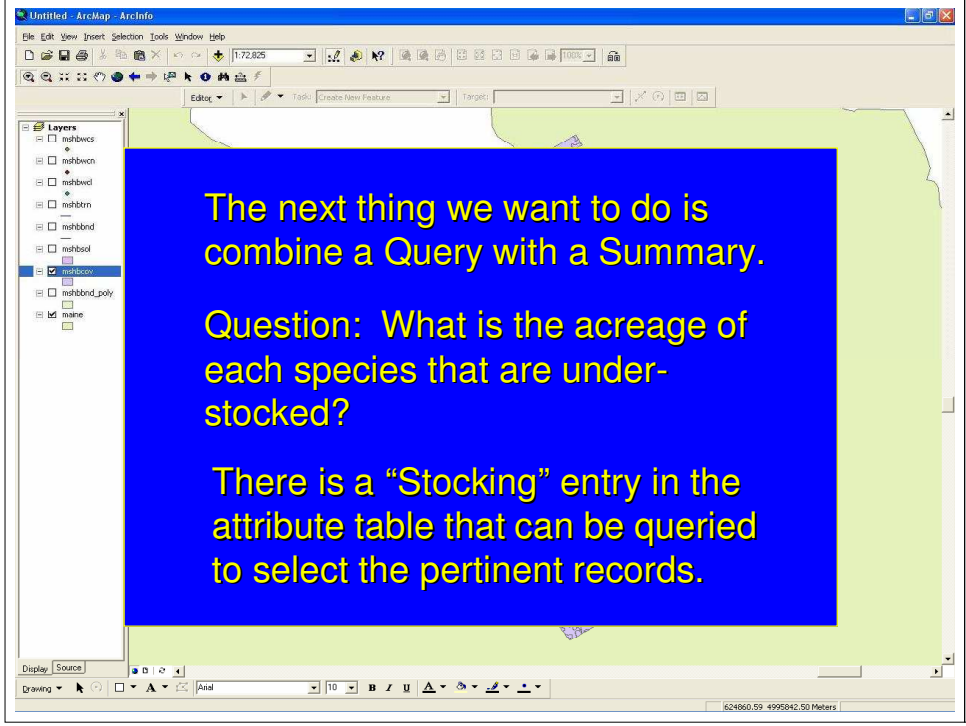

A wide range of statistical and summarization functions can be performed using the Graphical User Interface options in ArcMap.

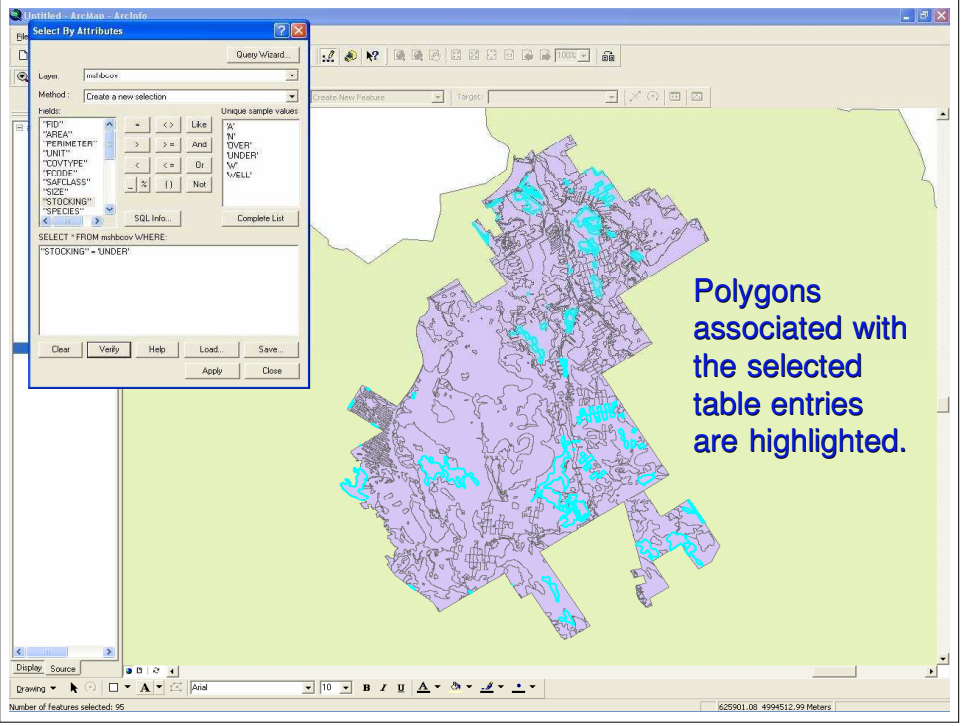

Click the SPECIES Header
Stepping through the details one time….

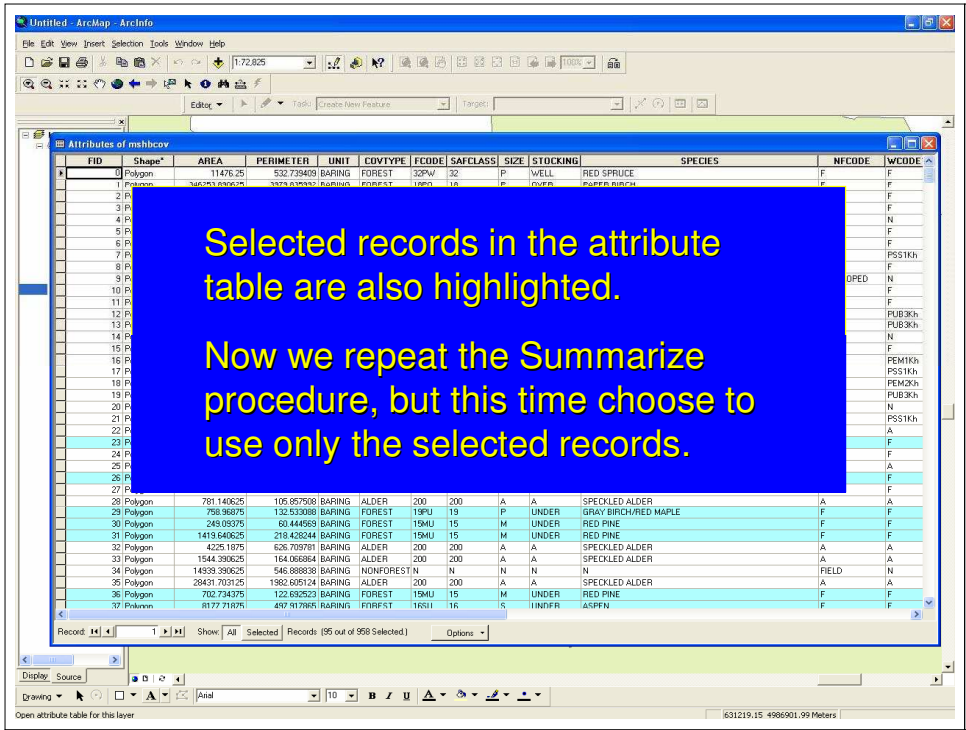

Choose to Summarize

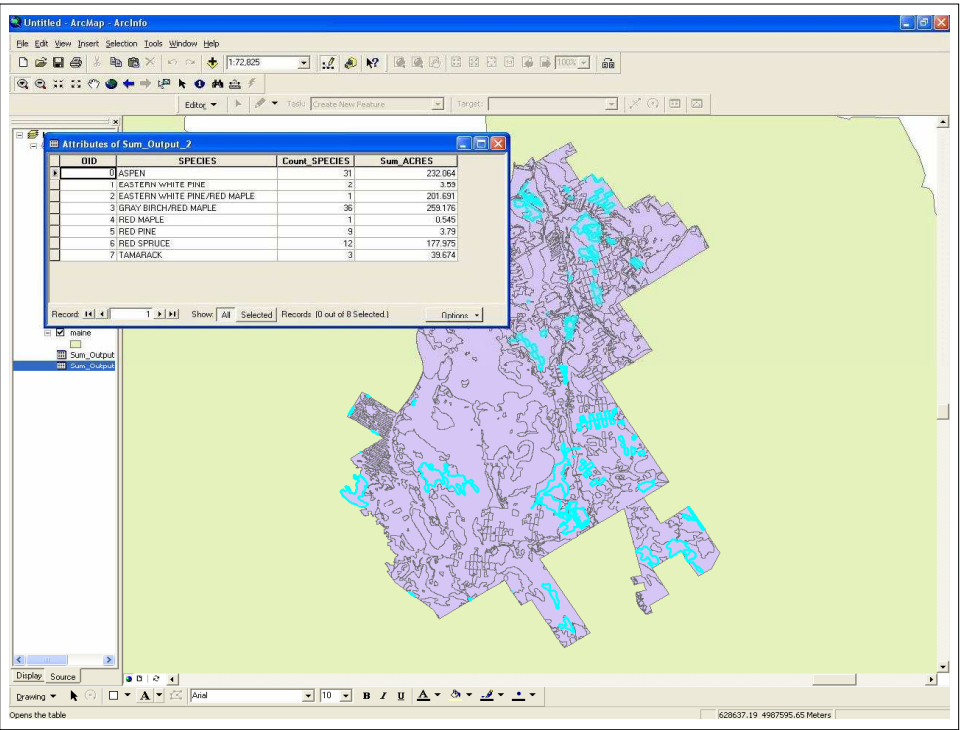

Untitled - ArcMap - ArcInfo
Attributes of mshbcov
Summarize
Summarize creates a new table containing one record for each unique value of the selected field, along with statistics summarizing any of the other fields.
1. Select a field to summarize:
SPECIES
2. Choose one or more summary statistics to be included in the output table:
ACRES
Minimum
Maximum
Average
Sum
Standard Deviation
Variance
AREA
CLASS
COVTYPE
3. Specify output table:
C:\SpatialWorkshop\MSHGIS\shapefiles\Sum_Output.db
More about Summarize...
OK
Cancel

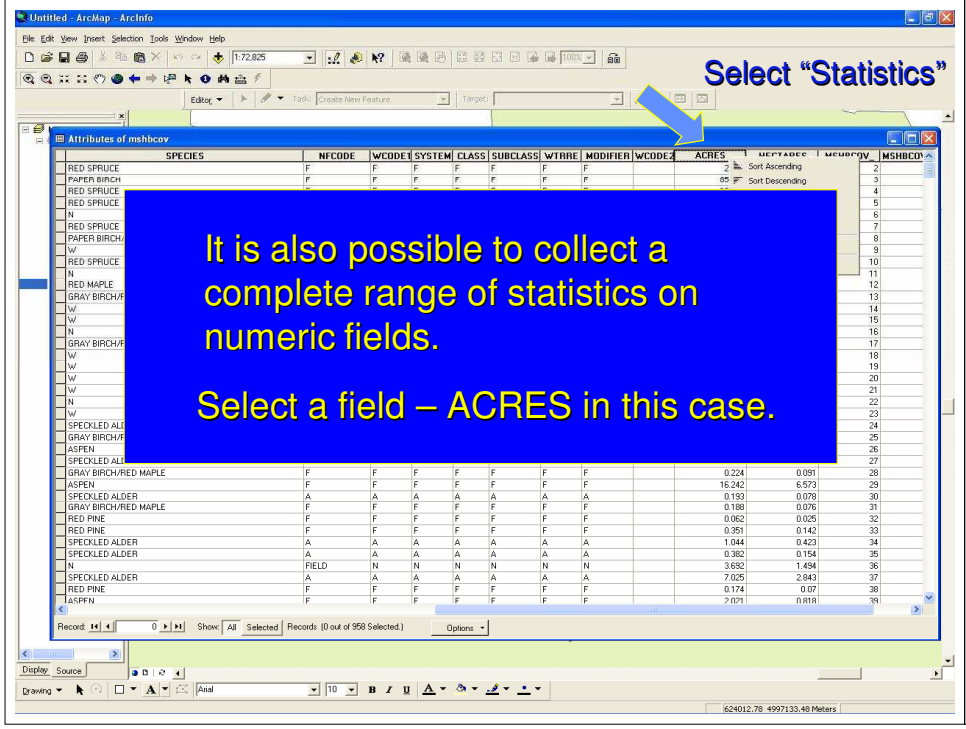

Untitled - ArcMap - ArcInfo
Attributes of Sum_Output
OID
SPECIES
Count_SPECIES
Sum_ACRES
1 ASPEN 238 5244.872
2 BLACK ASH/AMERICAN ELM/RED MAPLE 1 0.441
3 EASTERN WHITE PINE 25 737.998
4 EASTERN WHITE PINE/EASTERN HEMLO 2 25.846
5 EASTERN WHITE PINE/RED MAPLE 17 3573.897
6 GRAY BIRCH/RED MAPLE 126 948.679
7 N 64 360.408
8 NORTHERN RED OAK 2 74.754
9 NORTHERN WHITE CEDAR 7 46.296
10 PAPER BIRCH 2 89.579
11 PAPER BIRCH/RED SPRUCE 22 686.606
12 PIN CHERRY 3 15.283
13 RED MAPLE 17 405.195
14 RED PINE 15 79.035
15 RED SPRUCE 80 2141.666
16 SPECKLED ALDER 55 352.681
17 TAMARACK 16 187.238
18 W 265 2284.471

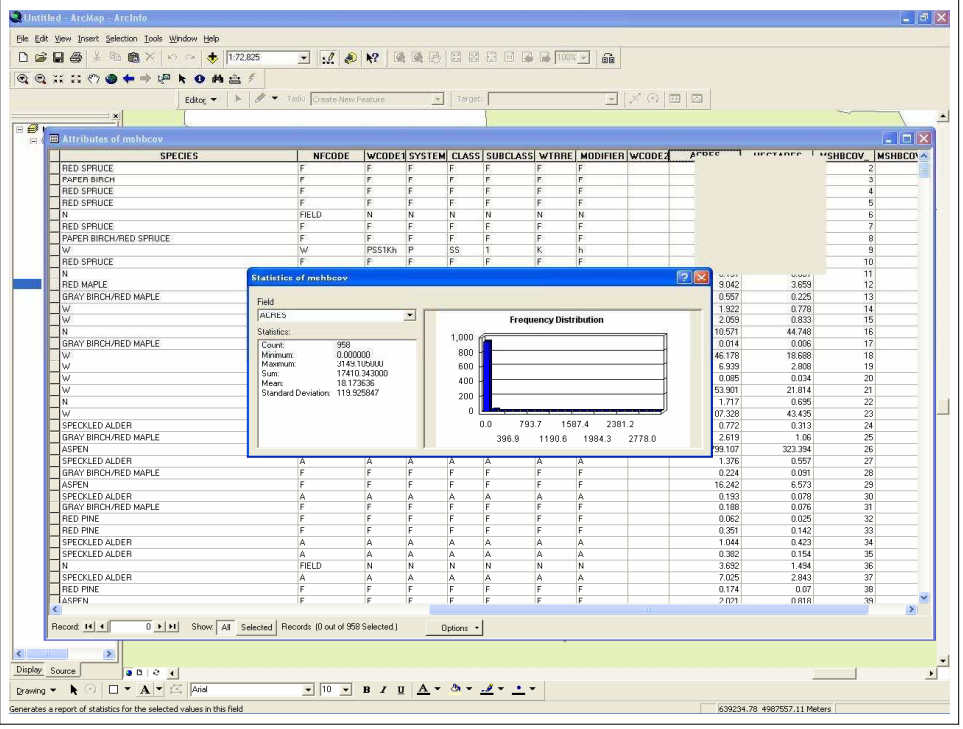

Untitled - ArcMap - ArcInfo
Layers
mshbcov
The next thing we want to do is combine a Query with a Summary.
Question: What is the acreage of each species that are under-stocked?
There is a “Stocking” entry in the attribute table that can be queried to select the pertinent records.


Untitled - ArcMap - ArcInfo
Select By Attributes
SELECT * FROM mshbcov WHERE:
"STOCKING" = 'UNDER'
Polygons associated with the selected table entries are highlighted.


Selected records in the attribute table are also highlighted.
Now we repeat the Summarize procedure, but this time choose to use only the selected records.




Select "Statistics"
It is also possible to collect a complete range of statistics on numeric fields.
Select a field – ACRES in this case.

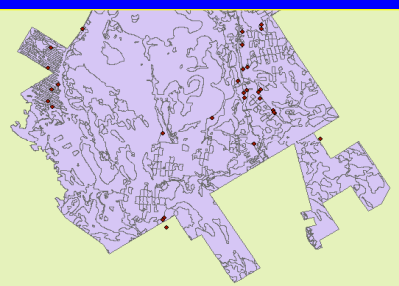


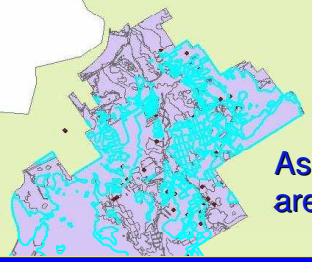
The next example is to find point features that are completely within polygon features.

Task: Find woodcock nest locations that are completely within aspen stands at Moosehorn NWR.

We can display the point theme showing woodcock nest locations.

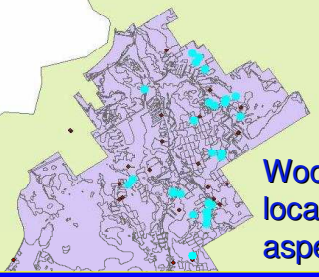
The next step is to select the aspen stands by the same method used to find the under-stocked polygons earlier.





Aspen stands are highlighted

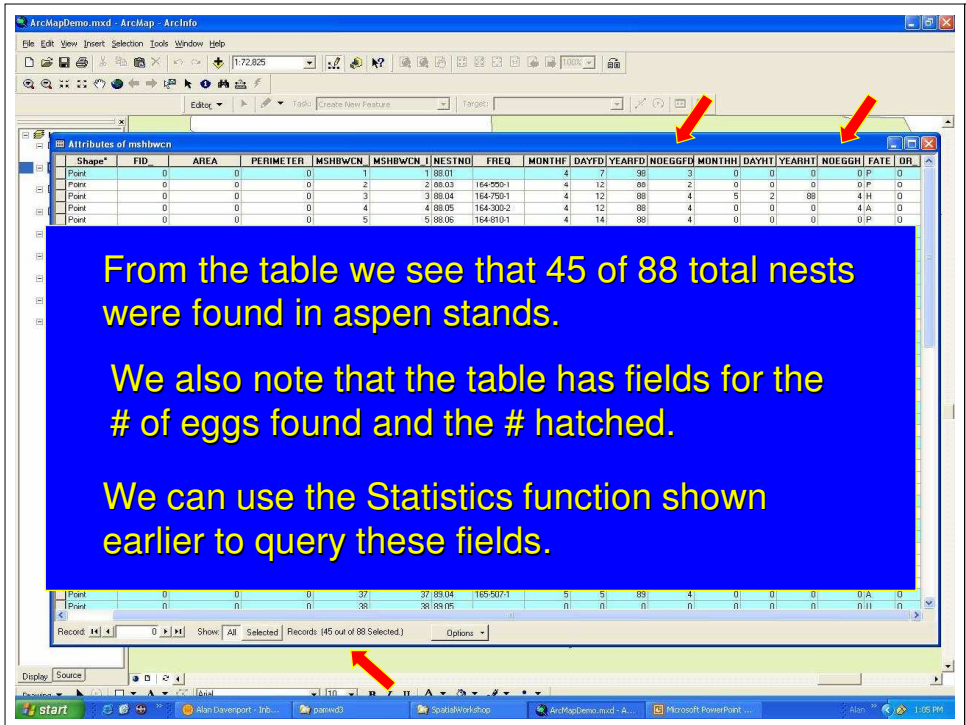
Using the query interface in ArcMap, we can ask for all woodcock nests that occur within the highlighted polygons (Aspen stands).

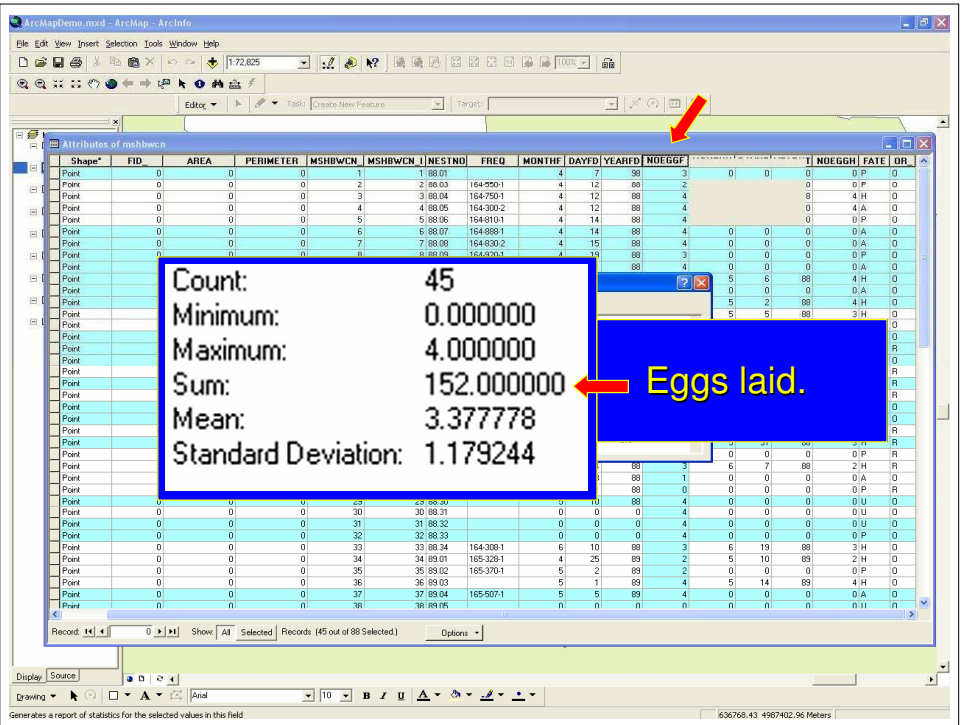


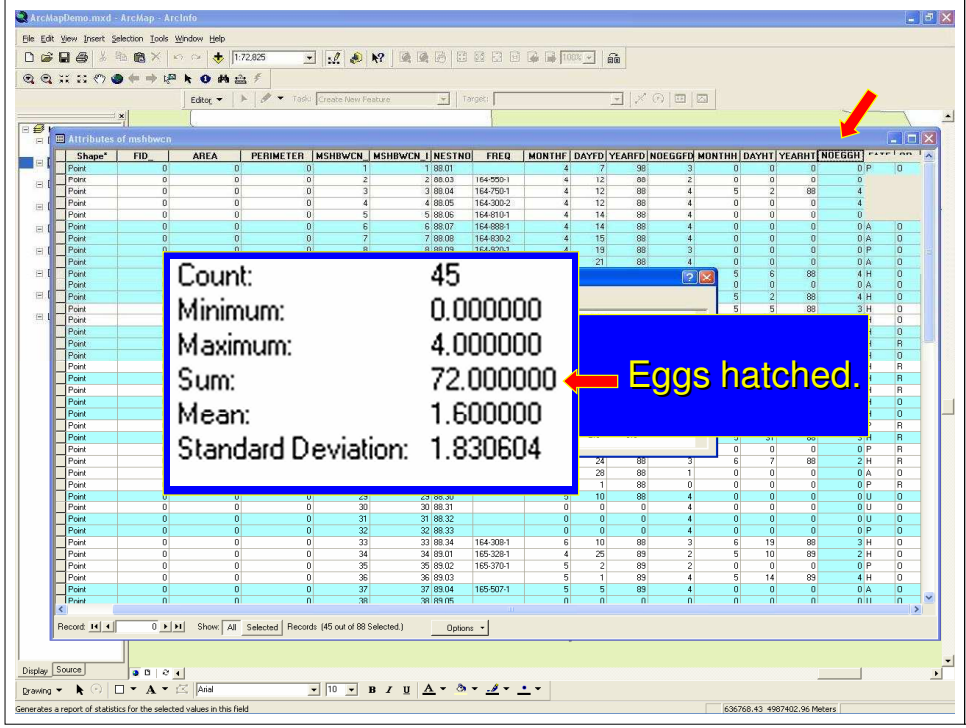
Woodcock Nest locations within aspen stands.

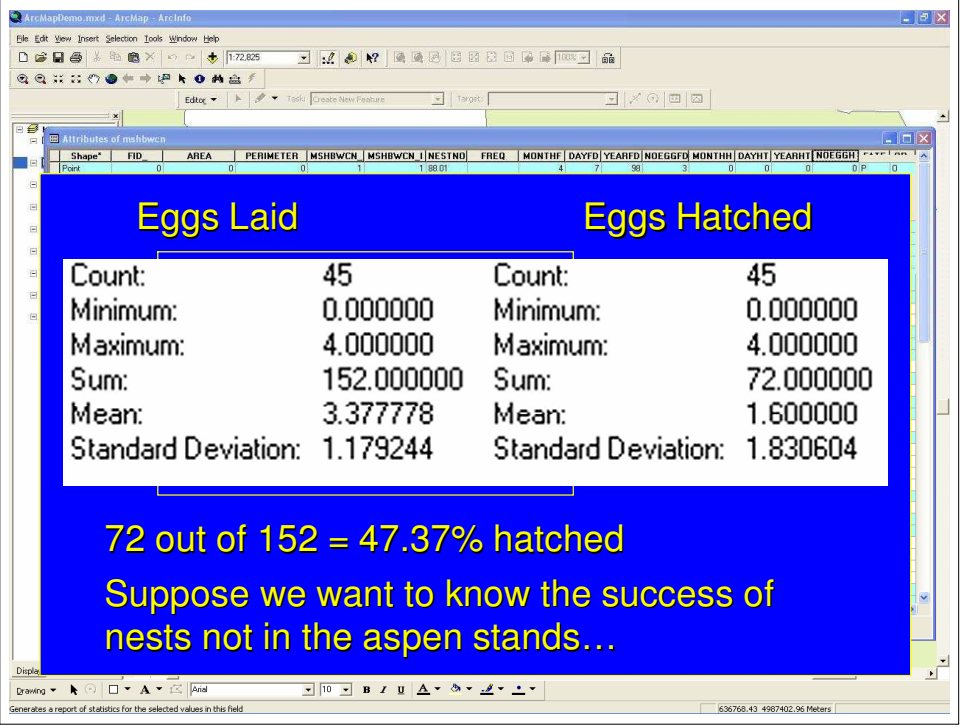
Suppose we wanted to know how many eggs hatched from these nests….

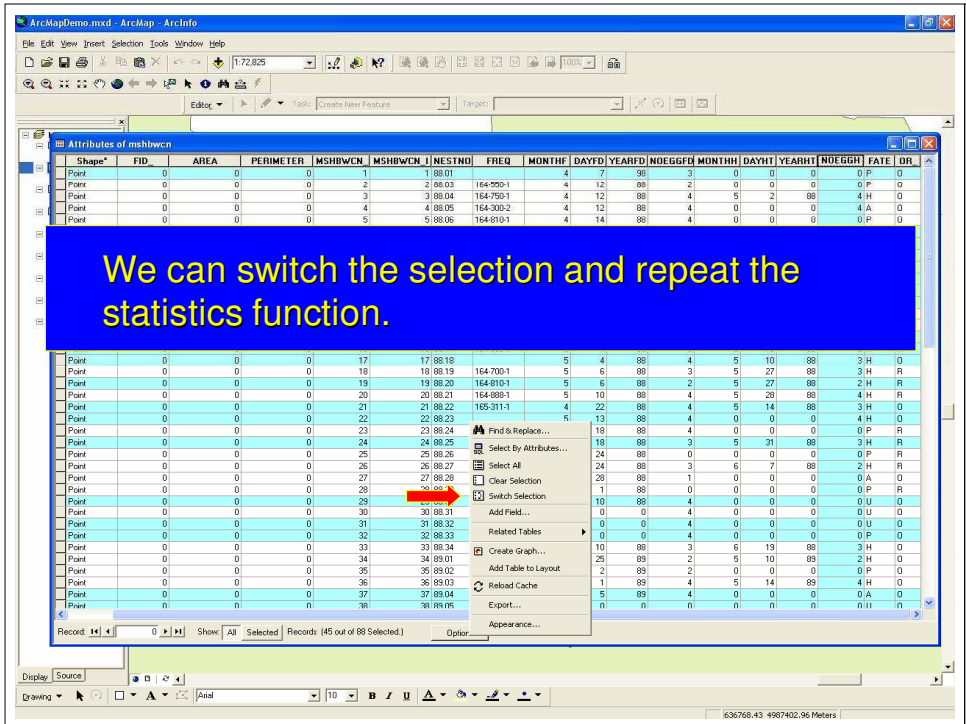
We would need to query the selected records in the nest attribute table.

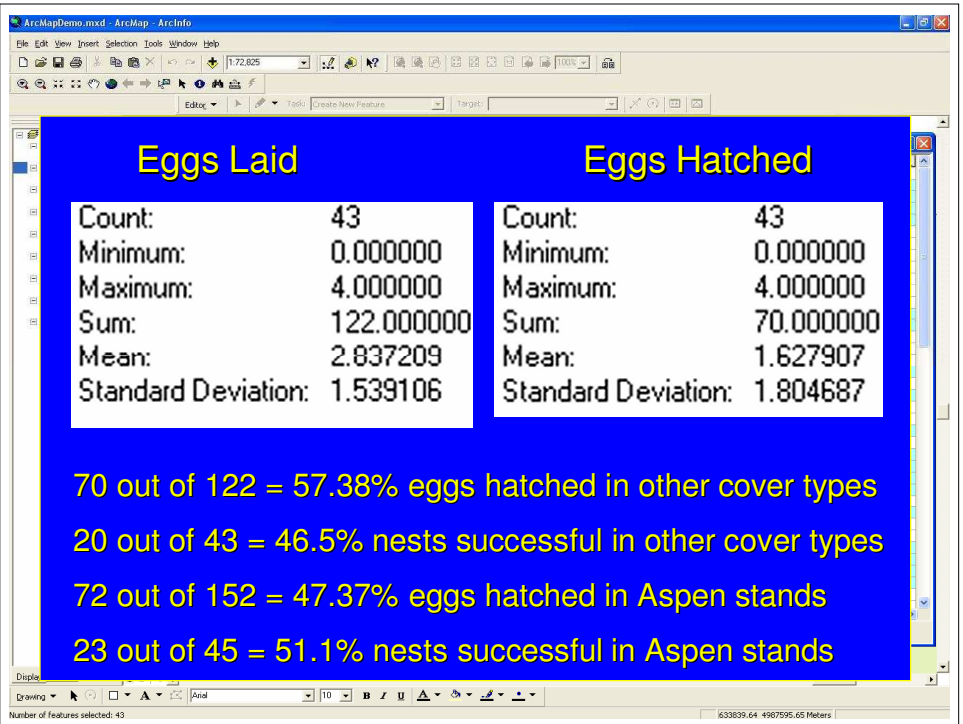

From the table we see that 45 of 88 total nests were found in aspen stands.
We also note that the table has fields for the # of eggs found and the # hatched.
We can use the Statistics function shown earlier to query these fields.

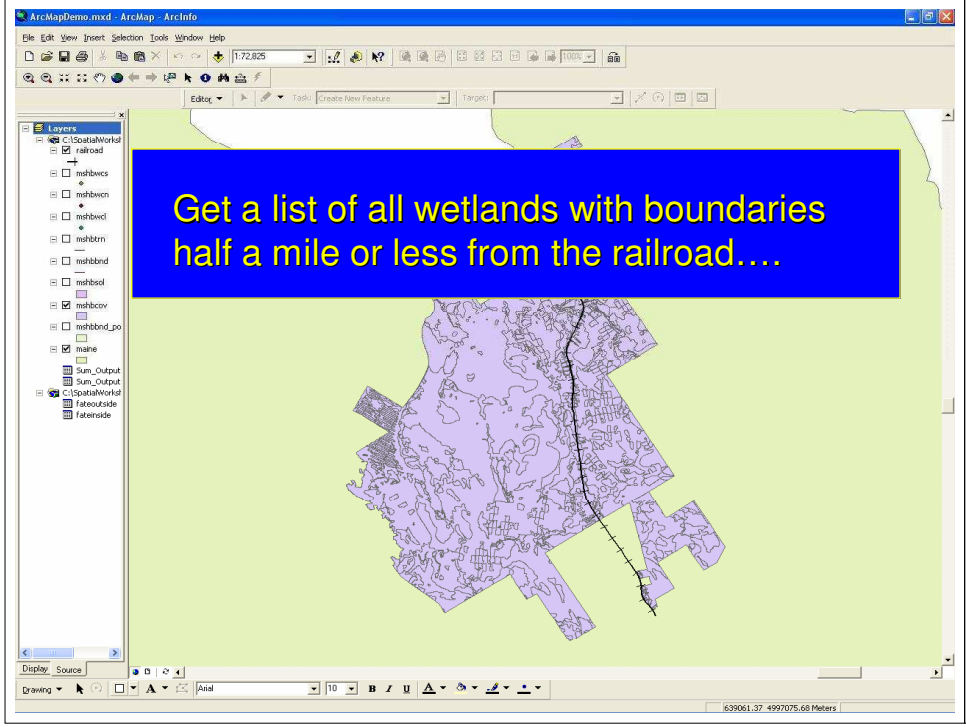

Count: 45
Minimum: 0.000000
Maximum: 4.000000
Sum: 152.000000
Mean: 3.377778
Standard Deviation: 1.179244
Eggs laid.

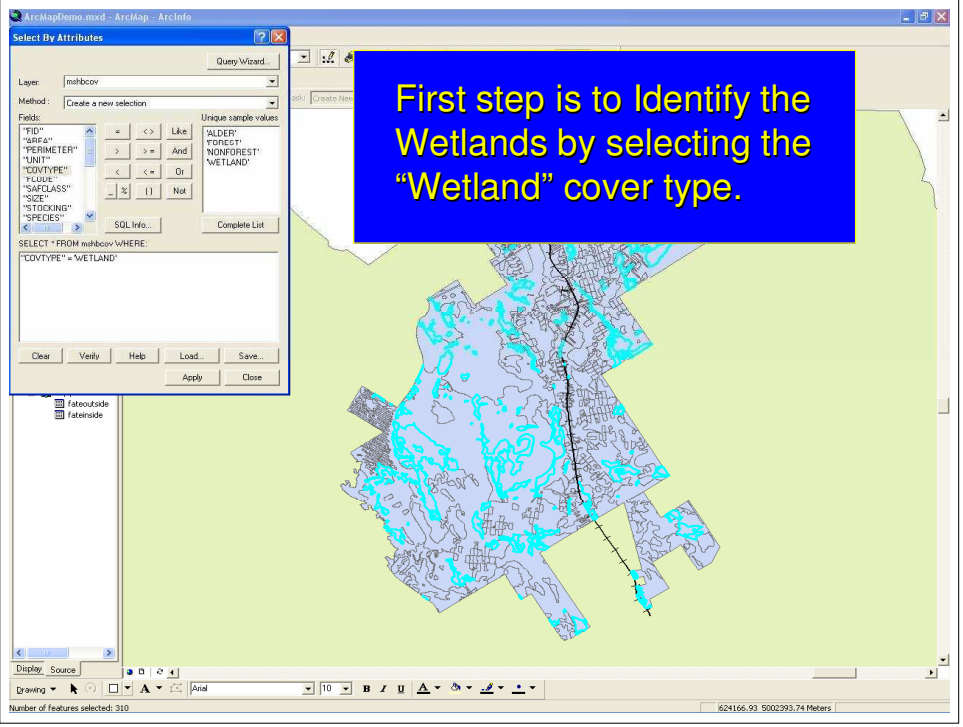

Count: 45
Minimum: 0.000000
Maximum: 4.000000
Sum: 72.000000
Mean: 1.600000
Standard Deviation: 1.830604
Eggs hatched.

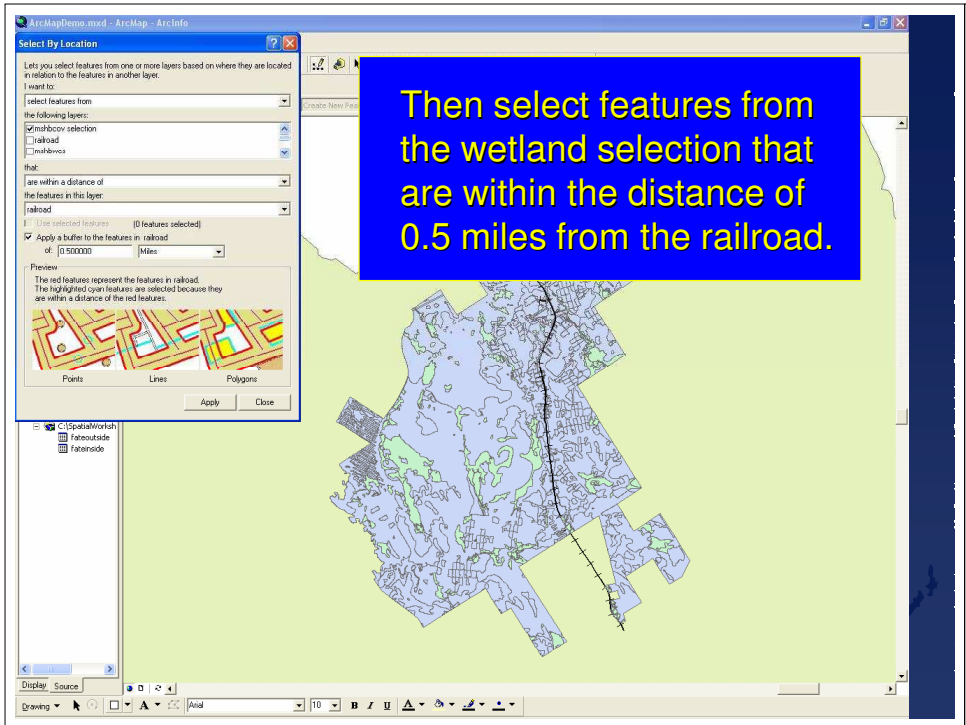

Eggs Laid
Count: 45
Minimum: 0.000000
Maximum: 4.000000
Sum: 152.000000
Mean: 3.377778
Standard Deviation: 1.179244
Eggs Hatched
Count: 45
Minimum: 0.000000
Maximum: 4.000000
Sum: 72.000000
Mean: 1.600000
Standard Deviation: 1.830604
72 out of 152 = 47.37% hatched
Suppose we want to know the success of nests not in the aspen stands…

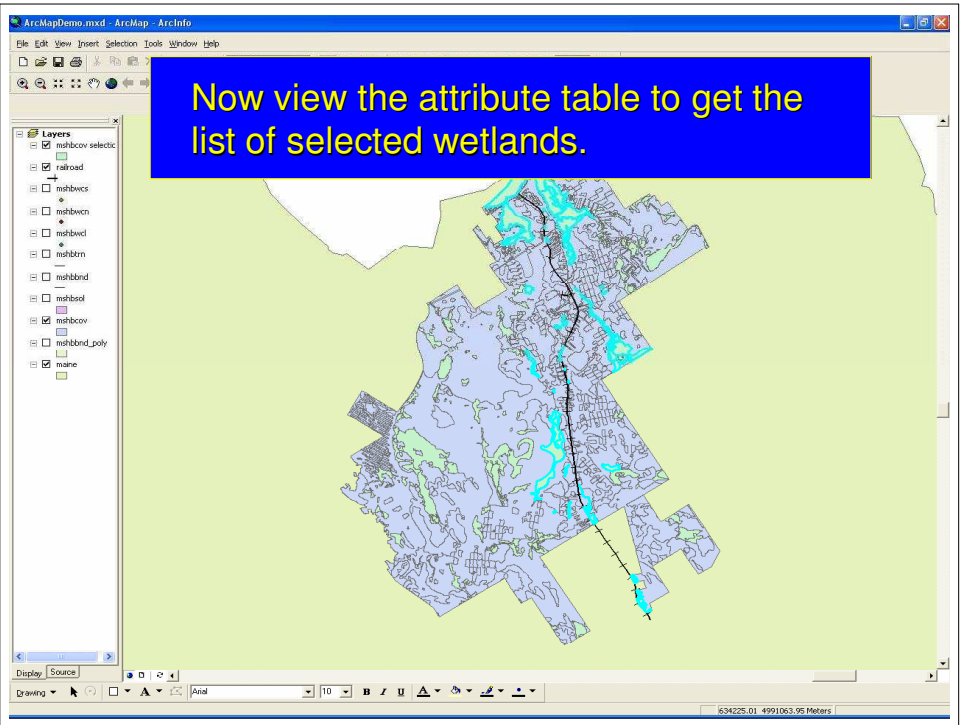

We can switch the selection and repeat the statistics function.

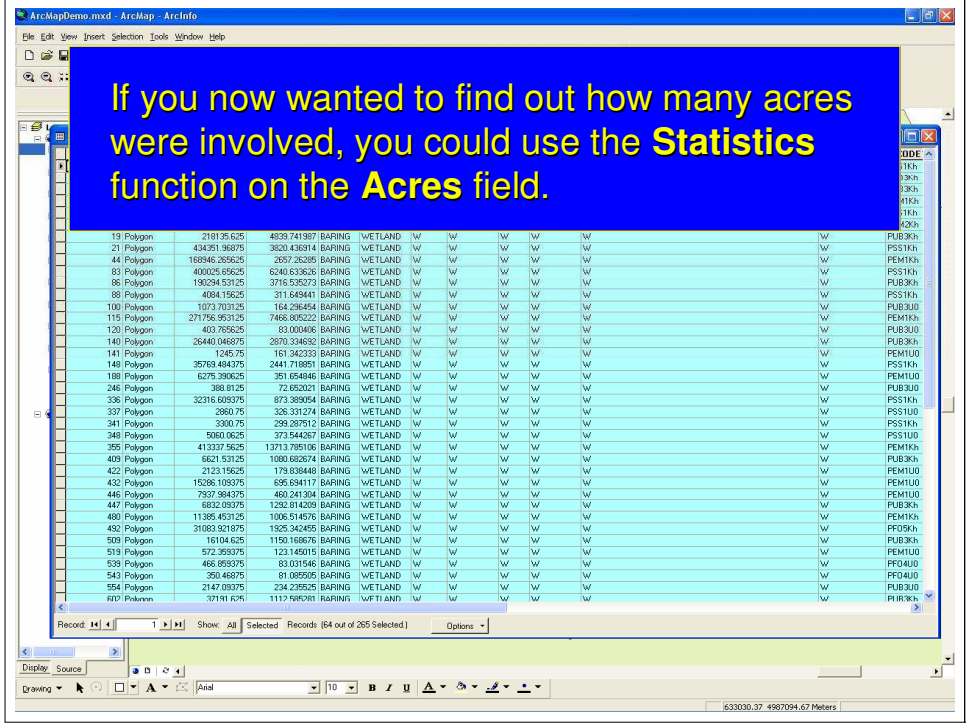

Eggs Laid
Count: 43
Minimum: 0.000000
Maximum: 4.000000
Sum: 122.000000
Mean: 2.837209
Standard Deviation: 1.539106
Eggs Hatched
Count: 43
Minimum: 0.000000
Maximum: 4.000000
Sum: 70.000000
Mean: 1.627907
Standard Deviation: 1.804687
70 out of 122 = 57.38% eggs hatched in other cover types
20 out of 43 = 46.5% nests successful in other cover types
72 out of 152 = 47.37% eggs hatched in Aspen stands
23 out of 45 = 51.1% nests successful in Aspen stands

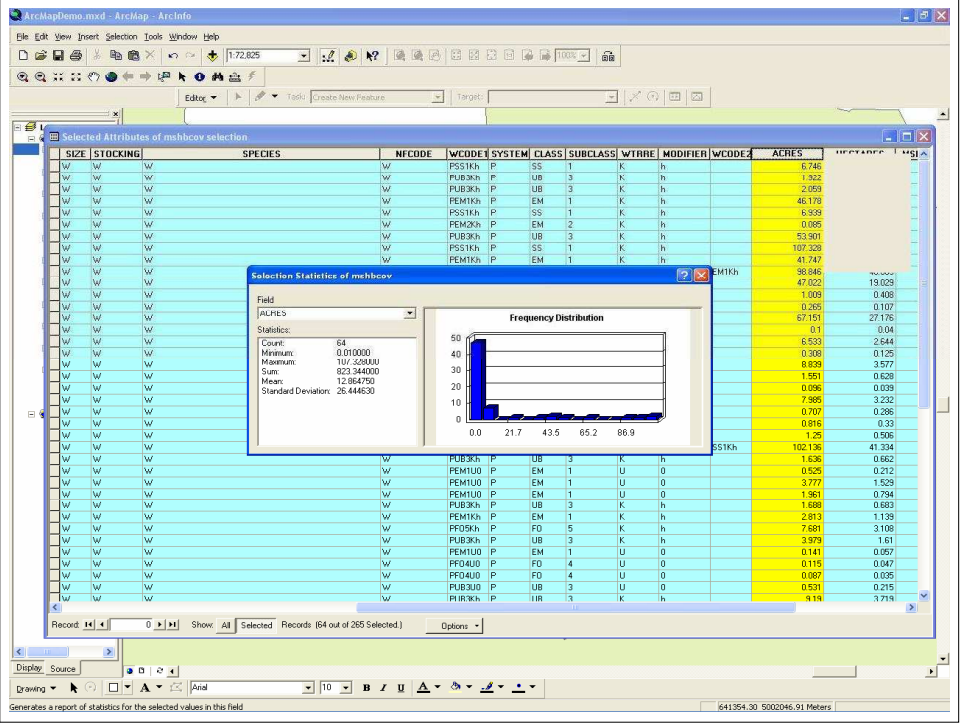

Get a list of all wetlands with boundaries half a mile or less from the railroad….


First step is to Identify the Wetlands by selecting the “Wetland” cover type.


Then select features from the wetland selection that are within the distance of 0.5 miles from the railroad.


Now view the attribute table to get the list of selected wetlands.


If you now wanted to find out how many acres were involved, you could use the **Statistics** function on the **Acres** field.

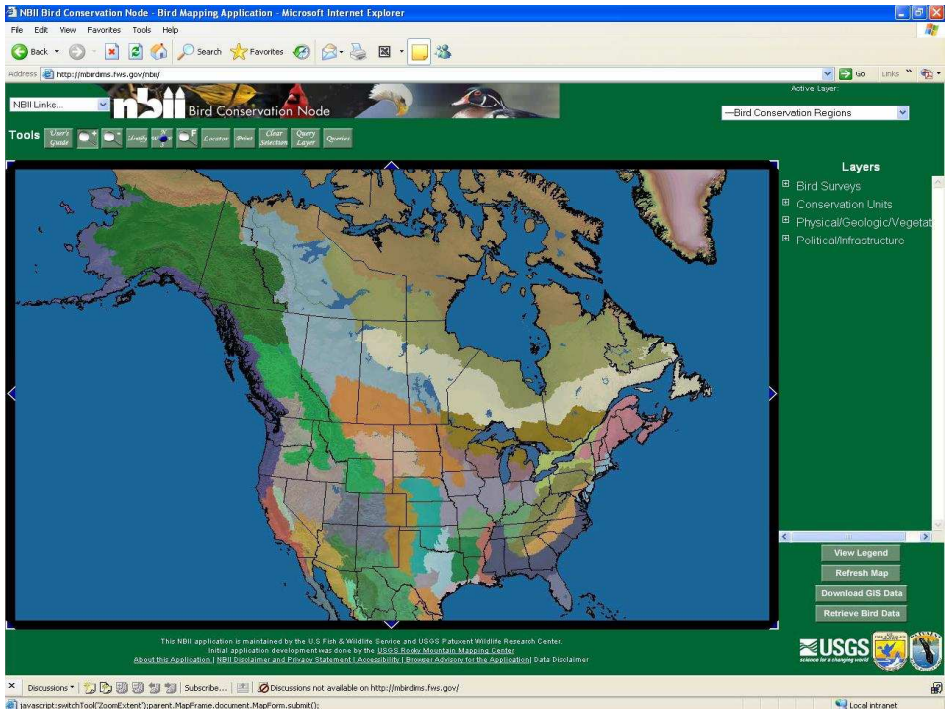


ArcIMS

- Internet Map Server
- Provides for viewing and manipulation of spatial data over the Internet.
- Our office hosts an ArcIMS application for the Bird Conservation Node of the NBII (National Biological Information Infrastructure)
- The application is reachable through the URL: http://mbirdims.fws.gov

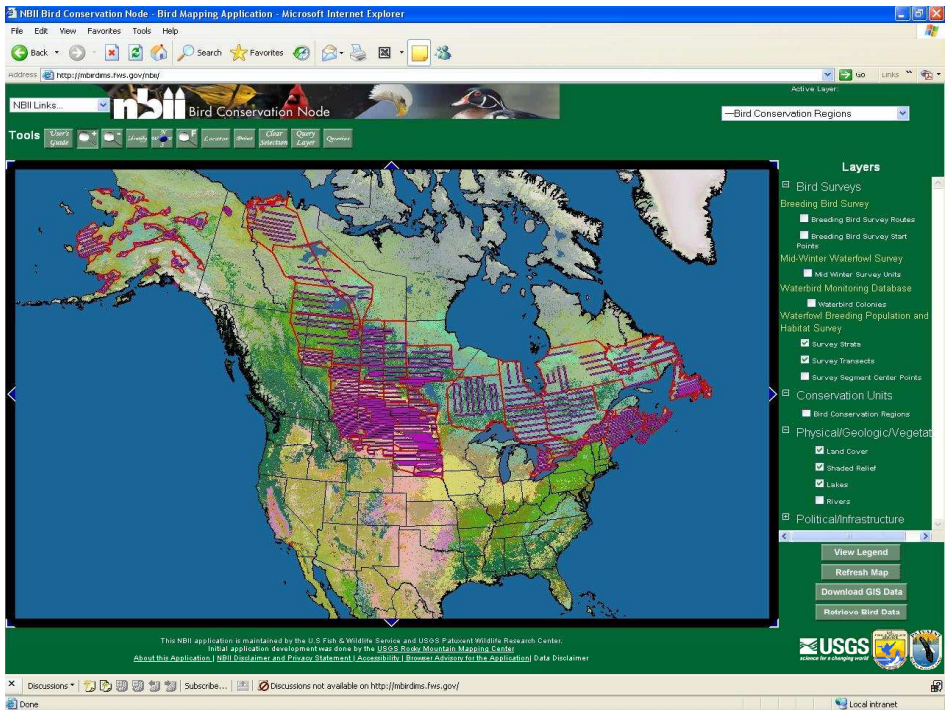
ArcIMS

- The next screen is the initial view presented when the web site is accessed. It shows most of North America. The shaded areas represent bird conservation regions. You can click on a "View Legend" button to view the key.



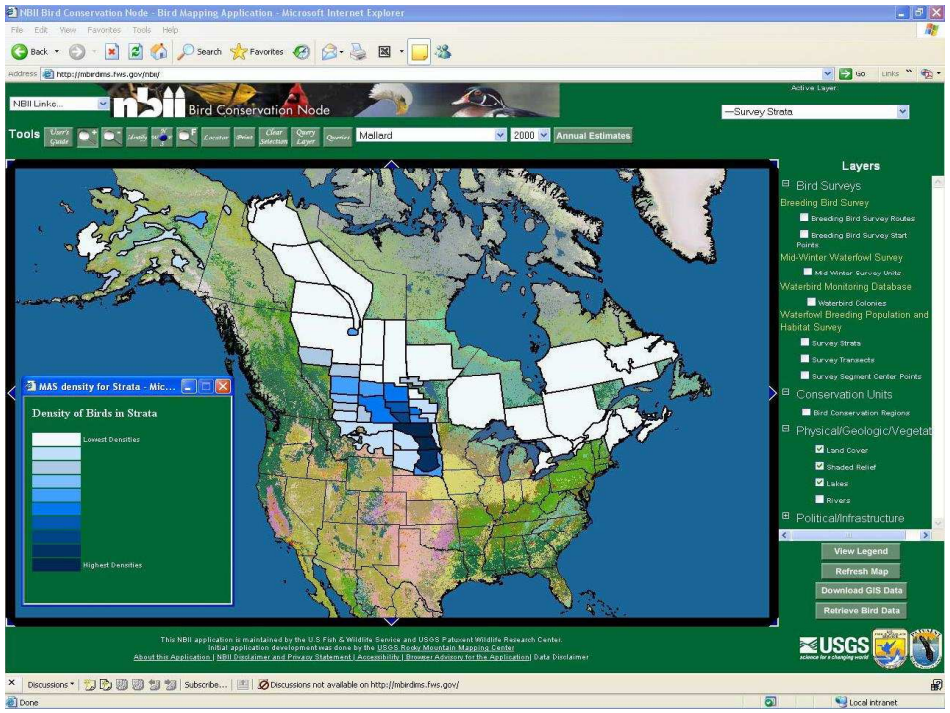
Aerial Surveys

- One of the major functions of our office is to sample the breeding grounds to estimate waterfowl populations annually.
- The next view shows the flight lines that are surveyed each year in May.



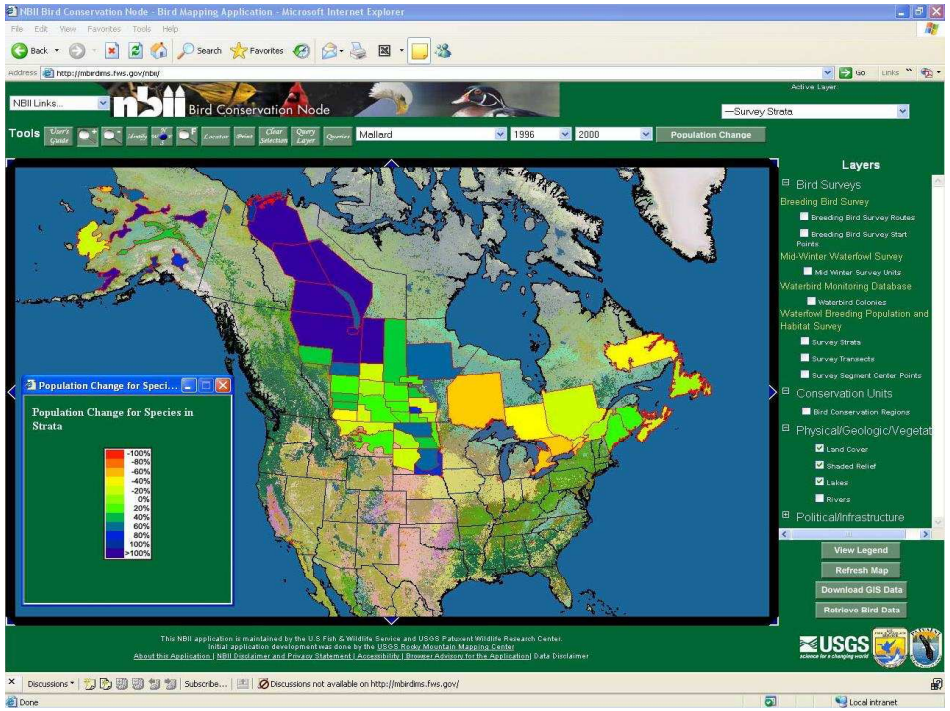
Aerial Surveys
(cont'd)

- An example of survey results is shown on the next slide.
- It shows the results of a query on Mallard Duck abundance by survey stratum.
- You may have noticed that the background has been replaced by a layer showing land cover types.



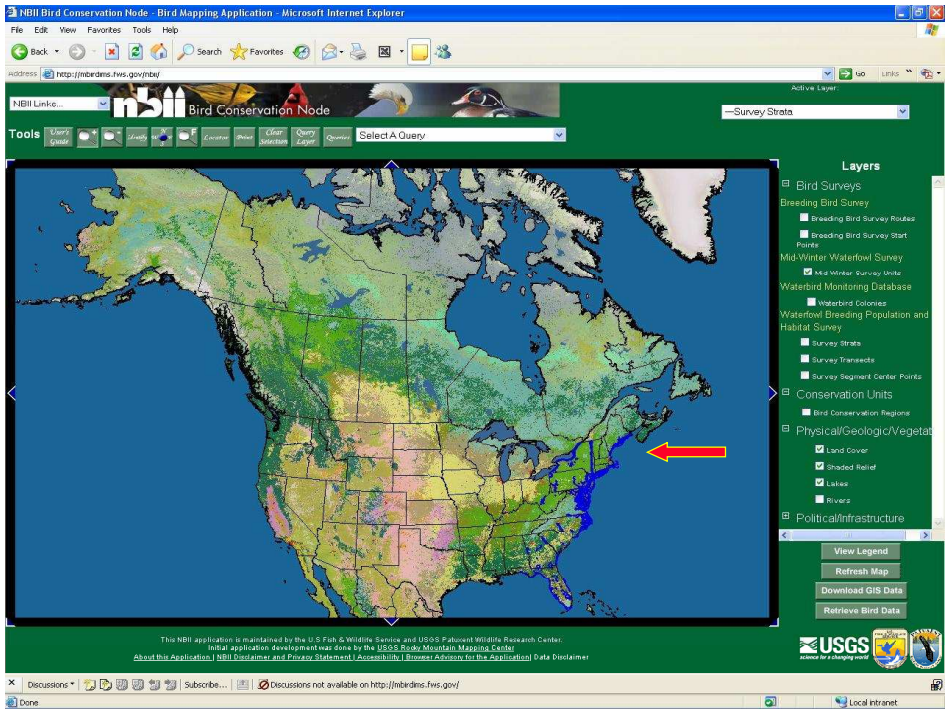
Aerial Surveys
(cont'd)

- The next example illustrates the changes in the population estimate for Mallards from 1995 to 2000.



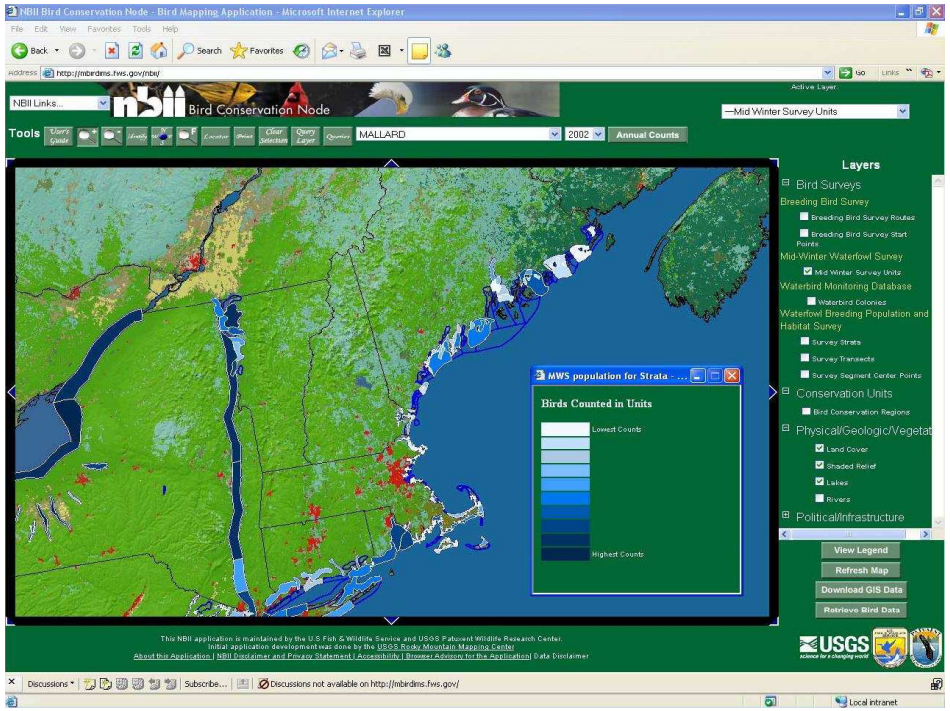
Aerial Surveys
(cont'd)

- In the next view the outlines for the Mid-Winter waterfowl survey zones have been displayed. This is a late December – January survey of wintering waterfowl, primarily along the Atlantic coast but covering inland waters of Atlantic coastal states.



Aerial Surveys
(cont'd)

- The next view shows counts of Mallards by survey unit for 2002.



Summary

This little demonstration hardly does justice to what can be done using a GIS to analyze data, but hopefully it has exposed the potentials.

But remember that you can't do anything unless you have the data necessary to answer the questions.

We'll finish up exploring a few methods for getting data into a GIS.

Polygons and Lines can be digitized from maps or other sources. Once initialized to known reference points, digitizing software automatically generates the correct geographic coordinates.

Points can be collected using GPS receivers.

Raster data is generated by photographs or satellite imagery.

Next up is a simple example of using satellite imagery as a background reference.

This is a satellite image, 978 pixels wide and 598 pixels high. The origin, pixel (0,0), is in the lower left corner.

In order to use this image effectively in a GIS, the coordinates of each pixel must be transformed to match the other layers being used.

Original Image from previous slide

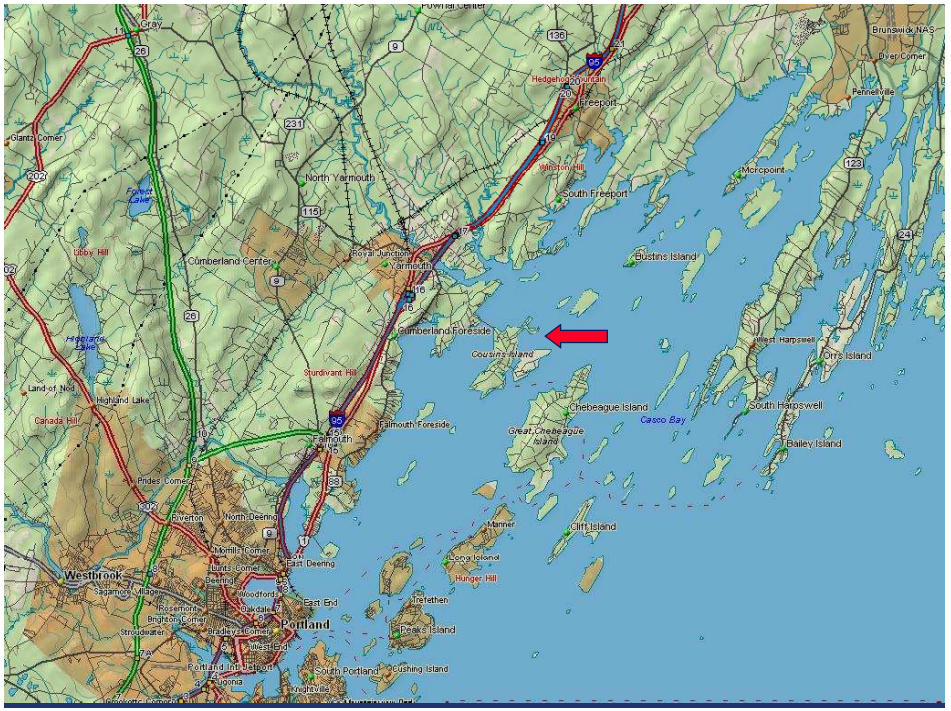
Transformed Image

Overlaid with Moosehorn cover type polygon outlines

Aerial Surveys

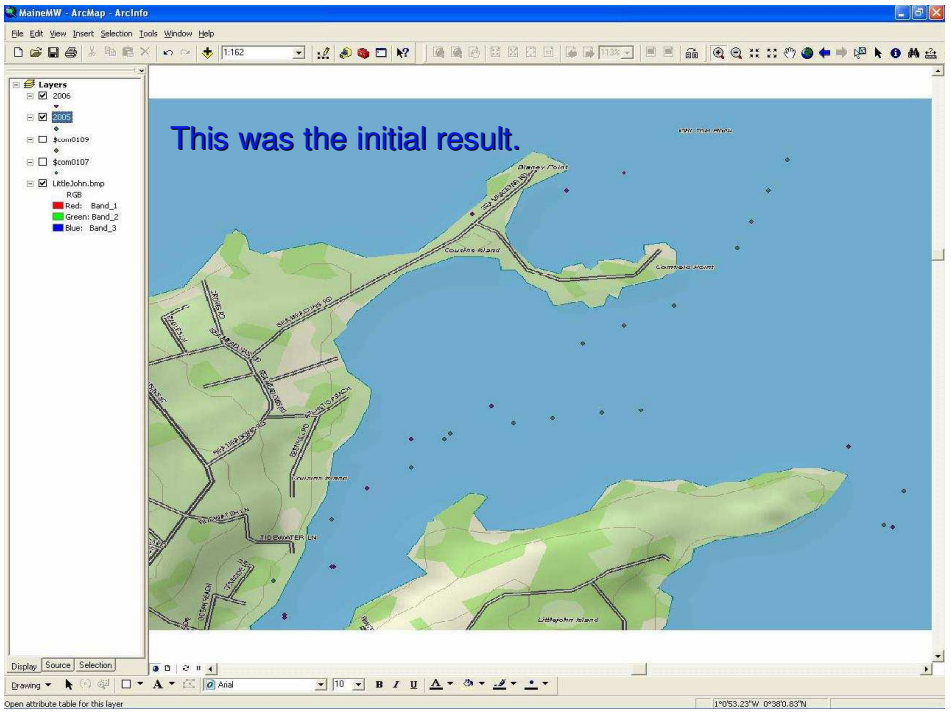
Aerial Surveys

- In January of this year I received a request through one of our pilot/biologists from a biologist in Maine who was interested in some specific data. There was concern about some development plans on Little John Island off the coast of Maine east of Yarmouth. He wondered if we could provide information about waterfowl species and counts in the area from the last two mid-winter waterfowl surveys.



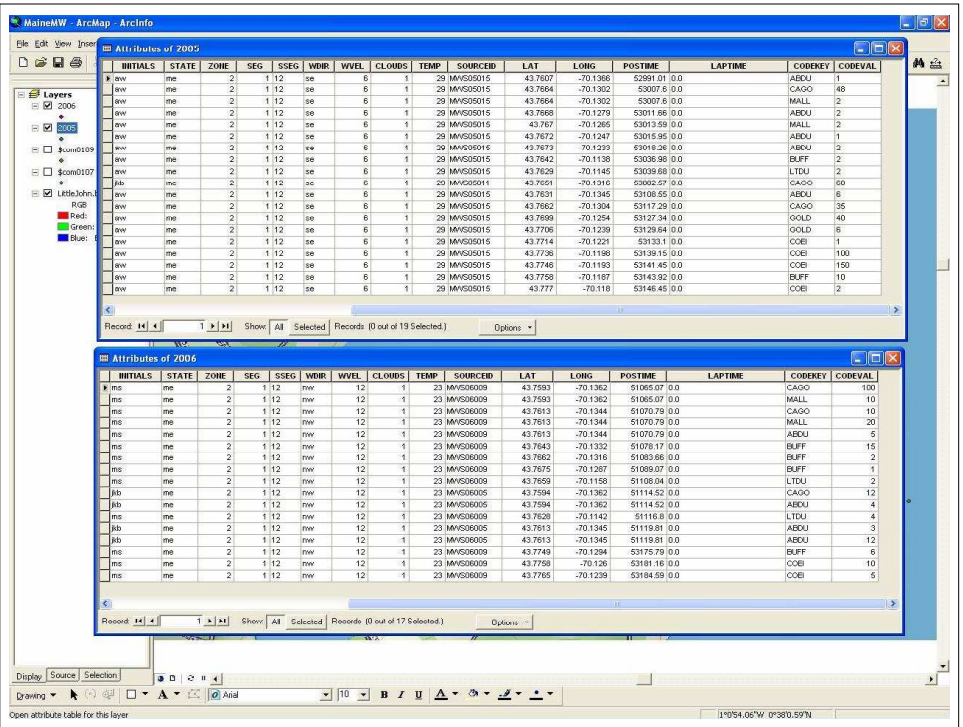
How to display "raw" survey data in a meaningful manner?

- The first thing was to create a base layer map to define the area of interest. Then the survey data points were plotted and those in the area of interest were selected using a "Select Features" tool which allowed me to "box" the area of interest and extract the attribute information for the points of interest.



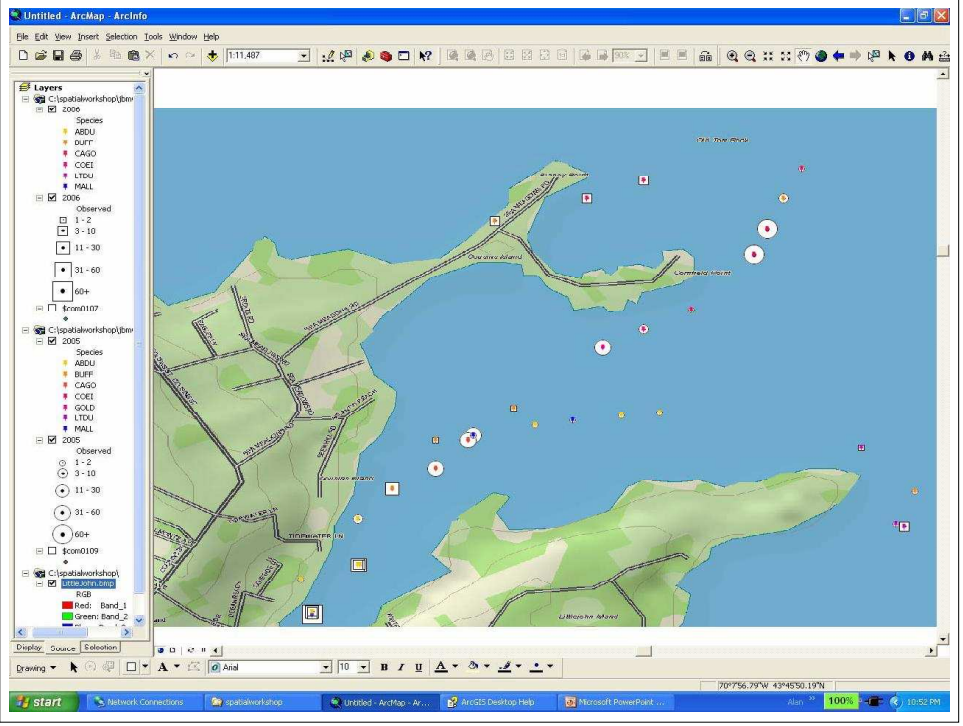
How to display "raw" survey data in a meaningful manner? (cont'd)

- There were not really many observations in the "target" area.
- Next we look at the attribute tables for the selected observations each year.



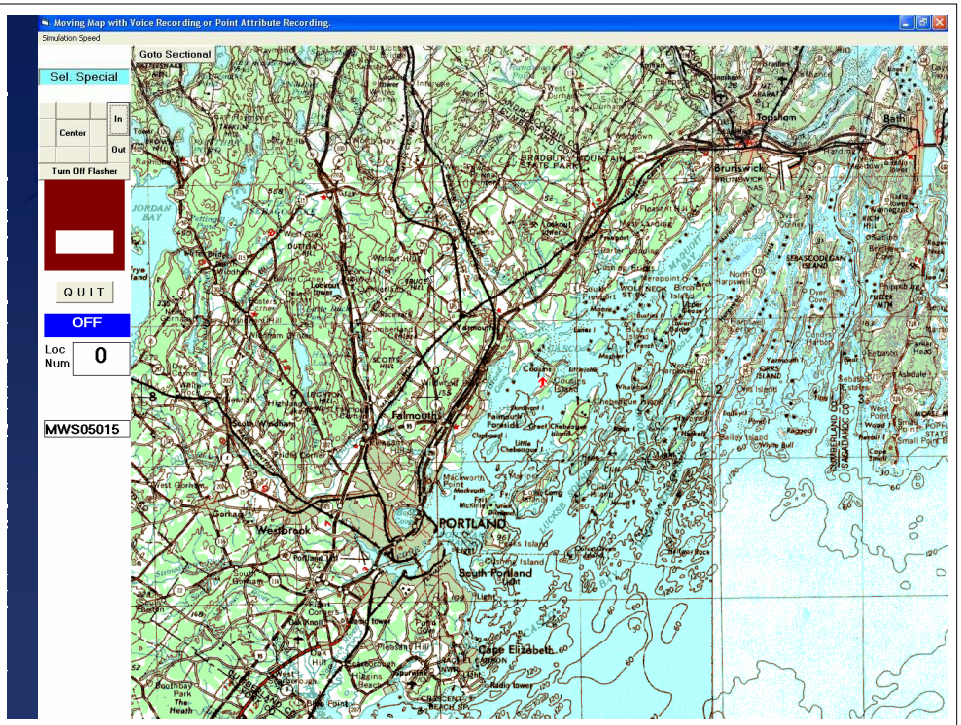
How to display "raw" survey data in a meaningful manner? (cont'd)

- Finally, symbology was selected to try and display the information in a more meaningful way than just as dots on a map.



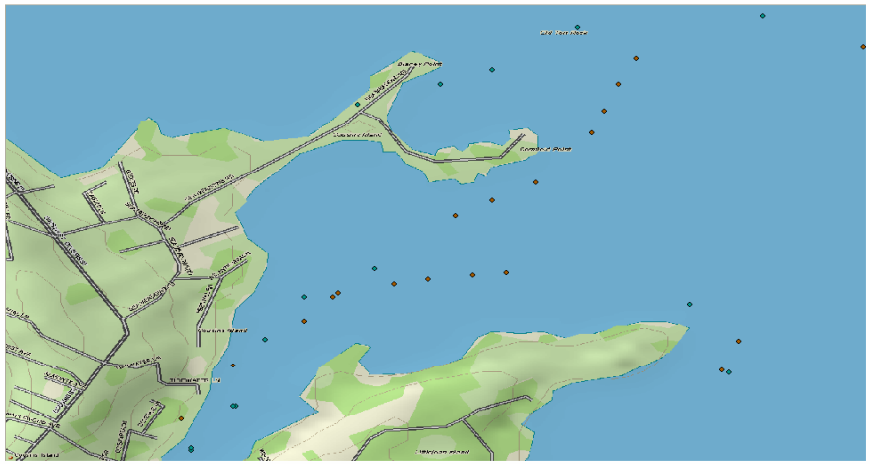
But how was this information actually collected?

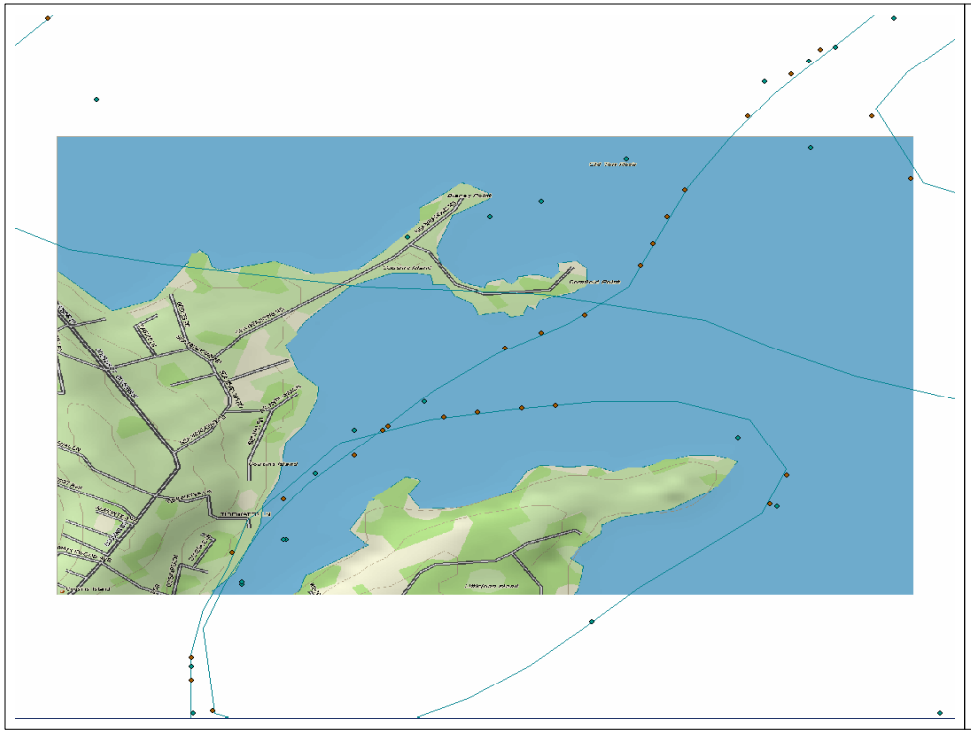
The next slide shows a simulation of part of the 2005 survey flight in the selected area.



Aerial Survey Data Collection

- Going back to the specific area of interest, we'll now see the data plots followed by an overlay of the flight path.






The End